



19th Meeting of the Data Quality Working Group

Proposal for approving amendments of S-100 Part 4C

Agenda Item 6.1A

DQWG-19, VTC Event, 25- 26 March 2024



IHO

BACKGROUND

International
Hydrographic
Organization

- Some texts in S-100 Part 4c refers to ISO 19138 Geographic Information - Data Quality Measurement, which has been withdrawn and revised by ISO 19157:2013 Geographic Information - Data Quality Standard.
- S-100 Part 4c needs to be revised to maintain consistency with ISO 19157 and IHO S-97.
- According to Action 18/17, the subWG comprised of Chair, NL and PRIMAR complete the review of S-100 Part 4C.



IHO

DISCUSSION/RECOMMENDATIONS

International
Hydrographic
Organization

It is recommended to:

- a) Replace the title of Appendix 4c-C with "Data Quality Metadata Attribute Definitions".

Note: S-100 is not only used in Hydrographic domain.



IHO

DISCUSSION/RECOMMENDATIONS

International
Hydrographic
Organization

It is recommended to:

b) Replace ISO 19138 with ISO 19157.

ISO Standards About us News Taking part Store Search

← TC ← ISO/TC 211

ISO/TS 19138:2006

Geographic information — Data quality measures

This standard has been revised by ISO 19157:2013



IHO

DISCUSSION/RECOMMENDATIONS

International
Hydrographic
Organization

It is recommended to:

- c) Replace "1Sigma", "2Sigma", "3Sigma", "4Sigma" and "5Sigma" with "68.3", "90", "95", "99" and "99.8", respectively.

linearMapAccuracy2Sigma90[0..1] : Real

Half length of the interval defined by an upper and lower limit in which the true value lies with probability 90%. [~~Adapted from ISO 19138~~] [Adapted from ISO 19157].

The Public Attribute is only used for vertical positional uncertainties.

linearMapAccuracy3Sigma95[0..1] : Real

Half length of the interval defined by an upper and lower limit in which the true value lies with probability 95%. [~~Adapted from ISO 19138~~] [Adapted from ISO 19157].

The Public Attribute is only used for vertical positional uncertainties.

linearMapAccuracy4Sigma99[0..1] : Real

Half length of the interval defined by an upper and lower limit in which the true value lies with probability 99%. [~~Adapted from ISO 19138~~] [Adapted from ISO 19157].

The Public Attribute is only used for vertical positional uncertainties.

**attributeValueUncertainty1Sigma 68.3[0..1] : Real**

This data quality measure indicates the attribute value of uncertainty where half the length of the interval defined by an upper and lower limit in which the true value for the quantitative attribute lies with a probability of 68.3%. ~~[Adapted from ISO 19138]~~ **[Adapted from ISO 19157]**.

attributeValueUncertainty2Sigma 90[0..1] : Real

This data quality measure indicates the attribute value of uncertainty where half the length of the interval defined by an upper and lower limit in which the true value for the quantitative attribute lies with a probability of 90%. ~~[Adapted from ISO 19138]~~ **[Adapted from ISO 19157]**.

attributeValueUncertainty3Sigma 95[0..1] : Real

This data quality measure indicates the attribute value of uncertainty where half the length of the interval defined by an upper and lower limit in which the true value for the quantitative attribute lies with a probability of 95%. ~~[Adapted from ISO 19138]~~ **[Adapted from ISO 19157]**.

attributeValueUncertainty4Sigma 99[0..1] : Real

This data quality measure indicates the attribute value of uncertainty where half the length of the interval defined by an upper and lower limit in which the true value for the quantitative attribute lies with a probability of 99%. ~~[Adapted from ISO 19138]~~ **[Adapted from ISO 19157]**.

attributeValueUncertainty5Sigma 99.8[0..1] : Real

This data quality measure indicates the attribute value of uncertainty where half the length of the interval defined by an upper and lower limit in which the true value for the quantitative attribute lies with a probability of 99.8%. ~~[Adapted from ISO 19138]~~ **[Adapted from ISO 19157]**.



d) Add the Public Attribute **BiasOfPositions** to **DQ_AbsoluteExternalPositionalAccuracy**.

In ISO 19157, the data quality measures for positional uncertainty in general of the data quality element **absolute or external accuracy** contains the **Bias of Positions**.

Table D.30 — Bias of positions

| Line | Component | Description |
|------|------------------|---|
| 1 | Name | bias of positions (1D, 2D and 3D) |
| 2 | Alias | - |
| 3 | Element name | absolute or external accuracy |
| 4 | Basic measure | not applicable |
| 5 | Definition | bias of the positions for a set of positions where the positional uncertainties are defined as the deviation between a measured position and what is considered as the corresponding true position |
| 6 | Description | <p>For a number of points (N), the measured positions are given as x_{mi}, y_{mi} and z_{mi} coordinates depending on the dimension in which the position of the point is measured. A corresponding set of coordinates, x_{ti}, y_{ti} and z_{ti}, are considered to represent the true positions. The deviation and biases are calculated as</p> <p>Single deviations:</p> $e_{xi} = x_{mi} - x_{ti}$ $e_{yi} = y_{mi} - y_{ti}$ $e_{zi} = z_{mi} - z_{ti}$ <p>Bias:</p> $a_x = \frac{\sum e_{xi}}{N_x}$ $a_y = \frac{\sum e_{yi}}{N_y}$ $a_z = \frac{\sum e_{zi}}{N_z}$ $a_p = \sqrt{a_x^2 + a_y^2}$ $a_{3D} = \sqrt{a_x^2 + a_y^2 + a_z^2}$ <p>A criterion for the establishing of correspondence should also be stated (e.g. allowing for correspondence to the closest position, correspondence on vertices or along lines). The criterion/criteria for finding the corresponding points shall be reported with the data quality evaluation result.</p> |
| 7 | Parameter | - |
| 8 | Value type | Measure |
| 9 | Value structure | - |
| 10 | Source reference | - |
| 11 | Example | - |
| 12 | Identifier | 128 |



IHO

DISCUSSION/RECOMMENDATIONS

International
Hydrographic
Organization

The revised content of S-100 Part 4C is as follows:

DQ_AbsoluteExternalPositionalAccuracy

Closeness of reported coordinative values to values accepted as or being true. [Per ISO 19115]

Public Attributes:

meanValuePositionalUncertainties[0..1] : Real

Mean value of the positional uncertainties for a set of positions where the positional uncertainties are defined as the distance between a measured position and what is considered as the corresponding true position. [~~Adapted from ISO 19138~~] [Adapted from ISO 19157].

BiasOfPositions[0..1] : Real

Bias of positions for a set of positions where the positional uncertainties are defined as the deviation between a measured position and what is considered as the corresponding true position. [Adapted from ISO 19157].



IHO

DISCUSSION/RECOMMENDATIONS

- e) Add texts to indicate which Public Attributes are only used for horizontal positional uncertainties and which Public Attributes are only used for vertical positional uncertainties in **DQ_AbsoluteExternalPositionalAccuracy**.

Sample revision of S-100 Part 4C is as follows:

linearErrorProbable[0..1] : Real

Half length of the interval defined by an upper and lower limit in which the true value lies with probability 50%. [~~Adapted from ISO 19138~~] [Adapted from ISO 19157].

The Public Attribute is only used for vertical positional uncertainties.

standardLinearError[0..1] : Real

Half length of the interval defined by an upper and lower limit in which the true value lies with probability 68.3%. [~~Adapted from ISO 19138~~] [Adapted from ISO 19157].

The Public Attribute is only used for vertical positional uncertainties.



IHO

DISCUSSION/RECOMMENDATIONS

- f) Correct the error in the example of the public attribute "rateOfExcessItems [0.. 1]" under **DQ_CompletenessCommissions**.

The revised content of S-100 Part 4C is as follows:

rateOfExcessItems[0..1] : Real

This data quality measure indicates the number of excess items in the dataset in relation to the number of items that should have been present. ~~[Adapted from ISO 19138]~~ [Adapted from ISO 19157].

This is a RATE which is a ratio, and is expressed as a REAL number representing the rational fraction corresponding to the numerator and denominator of the ratio.

For example, if there are 5 measured values and 4 valid values then the ratio is $5/4$ and the reported rate = 1.25 .



IHO

DISCUSSION/RECOMMENDATIONS

International
Hydrographic
Organization

- g) Correct the error in the example of the public attribute "rateOfMissingItems [0.. 1]" under **DQ_CompletenessOmissions**.

The revised content of S-100 Part 4C is as follows:

rateOfMissingItems[0..1] : Real

This data quality measure indicates the number of missing items in the dataset in relation to the number of items that should have been present. ~~—[Adapted from ISO 19138]~~ [Adapted from ISO 19157].

This is a RATE which is a ratio, and is expressed as a REAL number representing the rational fraction corresponding to the numerator and denominator of the ratio.

For example, if there are 3 measured values and 5 values are required the ratio is $\frac{3}{5}$ $\frac{2}{5}$ and the reported rate = 0.4.



h) Rename the Public Attribute **physicalStructureConflicts** of **DQ_FormatConsistency** as **physicalStructureConflictsNumber**, and add a new Public Attribute **physicalStructureConflicts** to **DQ_FormatConsistency** .

In ISO 19157, the data quality measures for **format consistency** contains **physical structure conflicts**, **physical structure conflicts number** and **physical structure conflict rate**.

D.3.3 Format consistency

The data quality measures for the data quality element format consistency are provided in [Tables D.19 to D.21](#).

Table D.19 — Physical structure conflicts

| Line | Component | Description |
|------|------------------|--|
| 1 | Name | physical structure conflicts |
| 2 | Alias | - |
| 3 | Element name | format consistency |
| 4 | Basic measure | error indicator |
| 5 | Definition | indication that items are stored in conflict with the physical structure of the data set |
| 6 | Description | - |
| 7 | Parameter | - |
| 8 | Value type | Boolean (true indicates physical structure conflict) |
| 9 | Value structure | - |
| 10 | Source reference | - |
| 11 | Example | True (data set is stored in wrong fileformat, shapefile instead of gml) |
| 12 | Identifier | 119 |

Table D.20 — Physical structure conflicts number

| Line | Component | Description |
|------|------------------|--|
| 1 | Name | number of physical structure conflicts |
| 2 | Alias | - |
| 3 | Element name | format consistency |
| 4 | Basic measure | error count |
| 5 | Definition | count of all items in the data set that are stored in conflict with the physical structure of the data set |
| 6 | Description | - |
| 7 | Parameter | - |
| 8 | Value type | Integer |
| 9 | Value structure | - |
| 10 | Source reference | - |



IHO

DISCUSSION/RECOMMENDATIONS

International
Hydrographic
Organization

The revised content of S-100 Part 4C is as follows:

DQ_FormatConsistency

Degree to which data is stored in accordance with the physical structure of the data set. [Per ISO 19115]

Public Attributes:

physicalStructureConflicts [0..1] : Integer

This data quality measure is an indication that items are stored in conflict with the physical structure of the dataset. [Adapted from ISO 19157].

physicalStructureConflictsNumber [0..1] : Integer

This data quality measure is a count of all items in the dataset that are stored in conflict with the physical structure of the dataset. [~~Adapted from ISO 19138~~] [Adapted from ISO 19157].

This is an integer count.



- i) Change the Public Attribute **temporalConsistencyStatement** of **DQ_TemporalConsistency** to **chronologicalOrder**.

In ISO 19157, the data quality measures for **temporal consistency** is **Chronological Order**.

D.5.2 Temporal consistency

One data quality measure for the data quality element temporal consistency is provided in [Table D.62](#).

Table D.62 — Chronological order

| Line | Component | Description |
|------|------------------|---|
| 1 | Name | chronological order |
| 2 | Alias | - |
| 3 | Element name | temporal consistency |
| 4 | Basic measure | error indicator |
| 5 | Definition | indication that an event is incorrectly ordered against the other events |
| 6 | Description | - |
| 7 | Parameter | - |
| 8 | Value type | Boolean (true indicates that the event is incorrectly ordered) |
| 9 | Value structure | - |
| 10 | Source reference | - |
| 11 | Example | True (5 historical events are present in the data set but are not ordered correctly). |
| 12 | Identifier | 159 |



IHO

DISCUSSION/RECOMMENDATIONS

International
Hydrographic
Organization

The revised content of S-100 Part 4C is as follows:

DQ_TemporalConsistency

Correctness of ordered events or sequences, if reported. [Per ISO 19115]

Public Attributes:

chronologicalOrder[0..1] : Boolean

This data quality measure indicates that an event is incorrectly ordered against the other events.

This is a Boolean where TRUE indicates that the event is incorrectly ordered. [Adapted from ISO 19157].

temporalConsistencyStatement[0..1] : CharacterString

~~This is a qualitative statement of the consistency of the time measurement.~~

~~There is no qualitative measure provided for this data quality sub-element. [Adapted from ISO 19138]~~



j) Add DQ_Aggregation.

S-97

ISO 19157

The components of Data Quality Measure can be divided into the following elements⁷:

1. Completeness
2. Logical Consistency
3. Positional Accuracy
4. Thematic Accuracy
5. Temporal Quality
6. Aggregation
7. Usability

D.7 Aggregation Measures

In a data product specification, several requirements are set up for a product to conform to the specification. The data quality measures for this element are provided in [Tables D.77](#) to [D.81](#).

Table D.77 — Data product specification passed

| Line | Component | Description |
|------|---------------|---|
| 1 | Name | data product specification passed |
| 2 | Alias | - |
| 3 | Element name | usability element |
| 4 | Basic measure | correctness indicator |
| 5 | Definition | indication that all requirements in the referred data product specification are fulfilled |
| 6 | Description | |



IHO

DISCUSSION/RECOMMENDATIONS

International
Hydrographic
Organization

The revised content of S-100 Part 4C is as follows:

DQ_Aggregation

Several requirements are set up for a product to conform to the specification. [Adapted from ISO 19157]

Public Attributes:

dataProductSpecificationPassed[O..1]: Boolean

This data quality measure indicates that all requirements in the referred data product specification are fulfilled. [Adapted from ISO 19157].

dataProductSpecificationFailCount[O..1]: Integer

This data quality measure indicates that the number of data product specification requirements that are not fulfilled by the current product/dataset. [Adapted from ISO 19157].

dataProductSpecificationPassCount[O..1]: Integer

This data quality measure indicates that the number of data product specification requirements that are fulfilled by the current product/dataset. [Adapted from ISO 19157].

dataProductSpecificationFailRate[O..1]: Real

This data quality measure indicates that the number of data product specification requirements that are not fulfilled by the current product/dataset in relation to the total number of data product specification requirements. [Adapted from ISO 19157].

dataProductSpecificationPassRate[O..1]: Real

This data quality measure indicates that the number of data product specification requirements that are fulfilled by the current product/dataset in relation to the total number of data product specification requirements. [Adapted from ISO 19157].



IHO

ACTION REQUIRED OF DQWG

International
Hydrographic
Organization

The DQWG is invited to:

- a) Note** the information provided;
- b) Approve** the amendment in Annex A as the new Appendix 4c-C of S-100;
- c) Take actions** to include Appendix 4c-C in the new edition of S-100.